DOCUMENT RESUME

ED 172 556 FL 010 374

AUTHOR Chiang, David L.
TITLE Predictors of Relative Clause Production.
PUB DATE [78]
NOTE 17p.; Paper presented at the Second Language Research Forum (Los Angeles, California); M.A. Thesis, University of Southern California

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Grammar; *Language Proficiency; *Language Research; *Predictor Variables; *Second Language Learning; *Sentence Structure; Student Characteristics; *Syntax; Verbal Tests
IDENTIFIERS *Relative Clauses

ABSTRACT

Predictors of relative clause production in adult second language acquisition were examined through a replication and extension of a study by Schachter (1974). The present study is different in that it was based on oral instead of written production and it was designed to examine avoidance with the consideration that avoidance presupposed choice. Two questions were asked of entering foreign students at the University of Southern California during an oral interview examination. One of the questions was meant to elicit a relative clause with a relativized subject, the other with a relativized non-subject. (Non-subject refers to direct object, indirect object, and oblique case NP's). Predictor variables were language background, overall target language proficiency, and the two input question types. A multiple regression analysis was performed. The results indicate that the order of importance of predictor variables in the determination of the frequency of the production of relative clauses is over-all target language proficiency, language background, and input question types. Although an avoidance strategy regarding relative clause production was found, there is a need to determine other contributing factors. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED172556

Predictors of Relative Clause Production

David L. Chiang

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
David L Chiang
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Department of Linguistics
American Language Institute
University of Southern California
Los Angeles, California 90007

FL010374

I. Setting and the Present Study:

The purpose of this study is to examine predictors of relative clause production in adult second language acquisition. Schachter (1974) found that Chinese and Japanese students in an English composition task, committed significantly fewer errors in using English relative clauses than did Persian and Arab students. But Schachter also pointed out that the number of relative clauses actually produced by Chinese and Japanese students was much lower than the other two groups and attributed this finding to native language interference, a fact predicted by contrastive analysis of the languages under study. Kleinmann, however, (1978,in Ritchie,ed.) argued that the observed 'relative nonuse' of linguistic feature/structure in Schachter's study was inconclusive in determining an avoidance strategy.

This present study, in general, is a replication and extension of Schachter's 1974 study. There are, however, two major differences: this study was (1) based on oral instead of written production and (2) was designed such that avoidance was examined with the consideration that avoidance presupposed choice. To properly define avoidance it is necessary to point out that being able to avoid also means being able to choose not to avoid (Kleinmann, 1977). By giving proper input stimuli to the subjects, this study provided them with the opportunity to 'choose' and was thus

able to study avoidance as defined above.

In this study, entering USC foreign students (n=83), during an oral interview examination, were asked two questions (cf. Appendix A.) as input stimuli. One of the questions was meant to elicit a relative clause with a relativized subject, the other with a relativized non-subject. (Non-subject, as used in this study, refers to direct object, indirect object and oblique case NP's.) The interviewers were instructed to ask the questions and then write down the students' answers to these two questions. The same set of questions was also given to native speakers of English (n=10) in a pilot test in order to confirm that the elicitation techniques were appropriate. It was found that answers with relative clauses were 'preferred' by native speakers. However, unlike studies of morpheme difficulty order, the 'use' of relative clause can not be said to be 'obligatory'.

In designing the study special consideration was given to the input question stimuli. It was argued by Coleman (1965) and re-advocated by Clark (1973) that input stimuli in language research may sometimes interact with the linguistic feature/structure under study. Thus the words used in this study were relatively easy so that they would not be a confounding factor in the production of the desired structure.

II. Statistical Analysis of Experimental Results:

Multiple regression analysis was used in this study. According to Kerlinger and Pedhazur (1973,3), multiple regression is " ... a method of analysing the collective and separate contribution of two or more independent variables to the variation of a dependent variable." In this study, the predictor variables are:

A. Language Background(LB): Following the findings in Schacter(1974),subject's native language was included as one predictor of the production of relative clauses. Five different languages were included and each of them was assigned a numerical coding value: Chinese(1); Japanese(2);Arabic(3);Persian(4) and Spanish(5). These numerical values served to indicate membership which ,in turn, was considered as an attribute (categorical) variable in the context of multiple regression analysis. Since this variable is 'categorical' in nature, it will be converted to dummy coding for regression analysis in order to avoid a biased result.

B. Overall Target Language Proficiency(OTLP): This variable indicated the students' English proficiency and was represented, also by numerical values, in accordance with the 'level' they were assigned to by the American Language Institute (090=0; 100=1; 101a=2; 101b=3; 102=4; released=5) on the basis of the overall ALI placement examination.

(Students assigned to 090 and 100 levels are not allowed to register for any other courses, whereas those at 101a, 101b, and 102 could register for one or two courses in their majors.)

C. Input Question Types(QT):Two different types of questions were coded in this fashion: 1=subject relativized; 2=non-subject relativized. Any student might have one type of question as the first one and the other type as the second or the other way around, but would never have two questions of the same type. (Since this variable was also 'categorical' in nature, it was, as the language background variable, converted into dummy vectors for unbiased analysis.)

The criterion variable in this study is the production or non-production of a relative clause in his/her response. This variable was coded as: 0=fail to produce; 1=produce preferred relative clause.

The coding of the variables in this study (except overall target language proficiency) involved nothing more than assigning attributes to different, well-defined, and multually exclusive categories.

All calculations were made with SAS (Statistical

Analysis System), a computational statistical package developed by Barr Goodnight, Sall and Helwig, SAS Institute Inc., N.C.

Table A presents the result of a Chi-Square test of the avoidance phenomenon.

---

insert table A. about here.

These Chi-square values show that there is a difference in terms of the use and non-use of relative clause construction in response to the questions.

Table B presents the intercorrelation matrix of Pearson correlation coefficients among all variables:

Insert Table B. about here

A multiple regression analysis (the maximum R-square improvement procedure, SAS User's Guide, Barr et al. 1976) was then performed on those two sets of data separately. Since the two different types of questions were randomly assigned to the students, these two sets of data can be regarded as the 'screening sample' and the 'calibration sample' as described by Kerlinger and Pedhazur. (Kerlinger & Pedhazur 1973, 284). By comparing the squared multiple correlation coefficients produced by the regression, it was found that the difference: $(R2^2=0.2206)-(R1^2=0.1816)=0.03$

taking into consideration that there might be a 'practice' effect on the second question, is regarded as a small amount of 'shrinkage'(ibid.) Therefore, the two sets of data were merged to provide a more stable prediction.

Table C presents the intercorrelation matrix of Pearson correlatin coefficients of the newly merged data set:

Insert Table C. about here.

Table D is a summary of the multiple regression analysis (Maximum R-square improvement procedure) performed on the merged data set:

Insert Table D. about here.

The results indicate that the order of importance of predictor variables in the determination of the frequency of the production of relative clauses is Over-all Target Language Proficiency; Language Background; and input Question Types. OTLP alone as predictor accounts for 10.421% of the variance in the production of RC's. The addition of LB, when entered into the regression equation next to OTLP, produces an increase of predicting power by (0.14646-0.10421)=4.23%. This amount of increase, after the significance test (as presented by Kerlinger and Pedhazur) is significant and meaningful. Likewise, the addition of QT

as the third predictor entered into the equation brings forth an increase of predicting power of (0.17136-0.14646)=3.59%, which again is significant and meaningful. Thus the production of English (non-center embedded) relative clauses, among other factors, is determined by OTLP; LB; and QT in that order.

III. Discussion:

This study has attempted to show that, in their speech, adult English learners resort to an avoidance strategy regarding relative clauses. This phenomenon was predicted by contrastive analysis and confirmed by Schachter and also found here . However, according to the result of this study, the avoidance strategy can not be attributed entirely to contrastive differences between native languages since language background (LGBK) was only the second best predictor of the production of relative clauses. The best predictor is overal language proficiency (OTLP), but even this predictor accounted for only about 11% of the variance. This indicates that other factors are probably active and should be explored.

table A.

Chi-Square Test of Avoidance

	LB	OTLP	QT
Avoidance: with or without RC	X^2=24.998 p < .0001	X^2=4.132 p < .05	X^2=11.202 p < .05

LB= Language Background

OTLP = Overall Target Language Proficiency

QT = Question Type

RC = Relative Clause

Table B.

Intercorrelation Matrix of Pearson Correlation Coefficients:

	LB	OTLP	QT(#1)	AN(#1)	QT(#2)
OTLP	-0.077				
QT(#1)	0.070	0.220			
AN(#1)	0.085	0.335	-0.158		
QT(#2)	-0.070	-0.220	-1.000	0.158	
AN(#2)	0.278	0.311	0.157	0.394	-0.157

LB=Language Background

OTLP=Overall Target Language Proficiency

QT(#1)=first question; AN(#1)=answer to the first question

QT(#2)=second question: AN(#2)=answer to the second question

Tabel C.

Correlation Matrix of Merged Data Set:

	LB	OTLP	QT
OTLP	-0.077		
QT	-0.000	-0.000	
AN	0.180	0.323	-0.158

OTLP=Overall Target Language Proficiency

QT=Question Type

AN=Answer

Table D.

Multiple Regression Analysis of Merged Data Set:

A. Predictor=OTLP; R^2=10.420% (Increment of R^2=0.00%)

Source	DF	Sum of Squares	ms=ss/df	F	Prob>F
Regression	1	4.2619	4.2619	19.08	0.0001
Error	164	36.6356	0.2233		
Total	165	40.8975			

B. Predictors=OTLP,LB; R^2=14.646% (Increment of R^2=4.23%)

Source	DF	Sum of Squares	ms=ss/df	F	Prob>F
Regression	2	5.9901	2.9950	13.99	0.0001
Error	163	34.9074	0.2141		
Total	165	40.8975			

C. Predictors=OTLP,LB,and QT; R^2=17.136% (Increment of R^2=3.59%)

Source	DF	Sum of Squares	ms=ss/df	F	Prob>F
Regression	3	7.0082	2.3360	11.17	0.0001
Error	162	33.8893	0.2091		
Total	165	40.8975			

Acknowledgements

The author would like to thank Dr. Edward Purcell for his valuable assistance with all the statistical analysis. Dr. Stephen Krashen, Dr. Eugene Briere, (Department of Linguistics, USC) and Dr. J. Schachter (ALI,USC) and Ms. Robin Scarcilla (Dept. of Linguistics, USC) also provided insightful opinions, comments and invaluable inspiration for this paper.

Bibliography

Barr Anthony J., et al (1976) A User's Guide to SAS, SAS Institute Inc., Raleigh, North Carolina.

Clark,H.H. (1973) "The language-as-fixed-effect fallacy: a critique of language statistics in psychological research," Journal of Verbal Learning and Verbal Behavior, 12, 335-359.

Coleman,E.B. (1964) "Generalizing to a language population," Psychological Reports, 14, 219-226.

Kerlger F.N. and E.J. Pedhazur (1973) Multiple Regression in Behavioral Research, Holt, Rinehart and Winston Inc. New York.

Kleinmann H.A. (1977) Avoidance behavior in adult second language performance. Language Learning, Vol.27, No.1. 93-107.

Kleinman, Howard H., (1978) "The strategy of avoidance in adult second language acquisition," in Ritchie William C. ed. Second Language Acquisition Research: Issues and Implications, Academic Press, New York.

Ritchie, William C. (1978) ed., Second Language Acquisition Research: Issues and Implications, Academic Press, New York.

Schachter,J. (1974) "An error in error analysis," Language Learning, 24,2. 205-213.

Appendix A. --Sample Questions--

Type 1 Questions: To elicit subject-relativized

Answers:

What is a teacher?

What is a musician?

What is a dentist?

Etc.

Type 2 Questions: To elicit object-relativized

Answers:

What is a typewriter?

What is a knife?

What is a bicycle?

Etc.